DOCUMENT RESUME

ED 130 528 FL 008 117

TITLE JURIVOC: Canadian Bilingual Legal Vocabulary. Presentation and Demonstration, COLING 76, June 28-July 2, Ottawa, Canada.
PUB DATE Jul 76
NOTE 16p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Bilingualism; *Computational Linguistics; *Computer Oriented Programs; *Computer Science; Dictionaries; Information Processing; *Laws; *Lexicography; Programing; Vocabulary; Word Lists
IDENTIFIERS Canada; Canadian Bilingual Legal Vocabulary; JURIVOC; *Legal Vocabulary

ABSTRACT

The use of computers in bilingual vocabulary processing and the task of compiling a Canadian bilingual legal dictionary are discussed. A model for automating lexicography and an analysis of the variables of bilingual communication are presented. It is emphasized that both human decision-making and computer manipulation are involved in this type of activity. The focus of this work is legal vocabulary which is written, formal, and of standard use in Canada since 1867. (SW)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130528

Vocabulaire Juridique Bilingue Canadien

JURIVOC

Canadian Bilingual Legal Vocabulary

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PRESENTATION and DEMONSTRATION

FL008117

COLING 76

June 28 – July 2

Ottawa Canada

AUTOMATED VOCABULARY PROCESSING OF BILINGUAL TEXTS

Any system of applied computational linguistics is symbiotic. It necessarily involves a certain mix of human decision-making and computer manipulation. System designers aiming at total automation soon identify the numerous human interventions required to help the machine. They also quickly isolate the work which a computer can do effectively. The particular mix upon which they decide will minimize time and cost and maximize accuracy. The variables of this design problem are individually discussed below.

The use of computers in bilingual vocabulary processing has permitted a modified approach to lexicography. Essentially, it has meant that the production of printed lexicons and dictionaries is secondary to the activity of teaching bilingual vocabulary to a machine which can print a dictionary at any time. This rationalization is of value because the gradual training of machines to recognize the different uses and translations of words means that will assume much of the future work of culling new texts and will indicate only those uses and translations which they do not understand and, in general, these will be the new ones.

Computer assistance to the lexicographical activity has permitted the undertaking of tasks hitherto practically unimaginable. Furthermore, it has re-introduced the possibility of the the goal of exhaustivity. In a period of a few minutes a computer can read, extract, sort and print the equivalent of 8 man-hours of work. This ratio will increase as lexicographers and linguists teach more to computers and become more sophisticated. This progressive automation develops as more decision-making power is identified, analysed, programmed and tuned for a computer. The visible tendency is to do more of higher quality at less expense. This positive trend has developed simultaneously with the proliferation of documentation and an accrued need for bilingual dictionaries as a means of disseminating linguistic conventions. Problem and solution always coexist.

Theoretical Framework for the Process

The model of lexicographical activity can be described as a series of states and events. The first state is the heterogeneous distribution of words and their related expressions linked to their correspondents. In other words, a specialized vocabulary in one language has been translated into another and has been recorded in a series of documents. The next state is the sorted list of all vocabulary in two languages structured by an identifiable order of precedence, that is, a file of the images of all words and expressions with their correspondents sorted alphabetically. The event which transforms the first state into the second is the analysis of all pairs of correspondents and their alphabetic ordering. This movement takes place on the whole set of words in a defined corpus of texts. The next event is a sequence of transformations each involving small sets of words and their related expressions. The third state is a normalized file of bilingual terminology synthesizing all distinguished uses and translations. While including some element of ponderation, the configuration is not necessarily prescriptive of use. The last state before the many regenerative cycles transform synergistically all the previous stages, is a bilingual thesaurus. This is prescriptive and is discovered by fitting words into hierarchic clusters of concepts which tend to describe the reality structure from which the vocabulary developed. The frame of analysis can be generic classes, special facets, principles, work systems or focii. The hierarchy may distinguish physical from metaphysical or try to reconcile the two in one description.

The above model of the lexicographical activity is dynamic and speaks in terms of states and transforming events. It has been derived from the approach of bionics which attempts to model technology after the behaviour of biological organisations. The human being senses the world and forms percepts or images of it. These images are stored with associated feelings that the individual experienced during perception. These sets of images are grouped into concepts which are labels for the set of experience images. The mind then further abstracts certain qualities of various groupings to form generalized principles. This hypothesis of human intelligence can now be referred to the model of lexicographical activity remembering that physical technology is but an extension, an amplifier, of human capacities.

The mere fact of recording symbols on paper, tape or disc extends the brain's memory function. The sorting of the elements by an order of precedence means that these images become easily accessible. The progressive normalization and structuring of individual elements or their clusters extends the unique function of the mind, discovering generalized principles. In fact, the lexicographical process is symbiotic. Humanity is collectively developing a mix which is tending to maximize the best of humans and machines. This proposition can be stated in terms of freedoms: computers liberate people from certain kinds of routinized work while allowing them freedom to exercise their characteristic intellectual functions for generalizing principles.

	STATE	EVENT	STATE	EVENT	STATE	EVENT	STATE
Model of Human Intelligence	Physical Environment & Human	Sensing	Percepts or Images in Brain	Review of Images	Concepts or Clusters of Images	Review of Concepts	Generalized Principles
Lexicographical Activity Model (artificial Intelligence)	Hetereogeneous Bilingual Vocabulary	Culling Texts	Vocabulary Listing Filed with Contexts	Different Uses Isolated	Lexicon	Normalization and Isomorphization	Thesaurus
Human Activities in Production Process		Entry, Linking, Group word families		Synthesis of Uses		Distinguishing and Defining	
Computer Programming	I/O Routines, Compare, Sorting		Storage & Retrieval System		Conversational Routines		Hierarchic Prompting System
Computer Activities		Reading, Sorting, Writing		Searching		Patterning	

A MODEL FOR AUTOMATING LEXICOGRAPHY

THE PROBLEM OF COMPUTER-AIDED BILINGUAL LEGAL LEXICOGRAPHY

INTRODUCTION

Lexicography is the writing or compiling of dictionaries.[1] A lexicographer studies how words are used; he collects as many examples as possible of a word's occurrence, analyses their differences and defines alternative meanings dependent upon context. While creating a definition or suggesting correct use, the lexicographer essentially crystallizes, from a large number of cases, the fundamental patterns of use of words and phrases.

The lexicographer could easily be accused of pretention because of the uncertainties of dealing not only with human communication but also with one or many languages. In addition he confronts the major problem of studying a large quantity of dispersed data. Several questions can be raised in response. Does not linguistic communication actually work despite its theoretical difficulties? And is not language a successful conventional system, however frail some might consider it? Is not transmission of meaning between linguistic groups feasible although sometimes imperfect? These questions provide the justification for the task of the lexicographer and recall the practical difficulties that are to be analysed in the following description of the JURIVOC systems. The computer is the powerful tool that has the potential to transform lexicography on an unprecedented scale. The discussion will successively treat communication, language, and bilingual environments to argue that while engaged in the traditionally arduous task of attempting to compile a bilingual legal dictionary there have developed some powerful practical computer aids for the lexicographer.

When talking of human communication, methodologists have frequently adopted the recent Shannon and Weaver model of information transmission.[2] It simply states that a message emanating from a transmitter is encoded, transported by some physical means using a source of noise, decoded and accepted by a receiver. While this model was developed for studying the transmission of messages along telephone wires it can serve as a starting framework for an analysis of bilingual linguistic communication.

Two humans can be substituted for the transmitter and receiver and the variables of the process of bilingual linguistic communication examined. The

verbal content of the message to be communicated is presumably known to the individual who desires to transmit some information. In fact, the knowledge of the verbal message can be more or less precise or vary when it is an emotional feeling or map knowledge. Furthermore, the originator might be incapable of communicating the message to himself hence practically precluding the possibility of communication with others. The first variable is therefore the quality of the knowledge to be communicated, its type and its clarity.

One can equally distinguish a variable related to the quality of knowledge of the receiving individual. An example is the complete lack of comprehension due to inexperience despite the apparent understanding of the code, or words, used. An illustration of this case might be an expert trying to inform a novice of a specialized concept in his field, or again an adult gourmet extolling the joys of good eating to a child. In both instances the receivers of the messages just do not know what the transmitters are saying. Thus a second variable is the spectrum of knowledge held by the receiver.

The intervening variables; encoding, decoding and code adequacy must now be examined. To explain the use of language as a coding method the classic linguistic triangle of sign, referent and reference shall be used.[3] Although many versions of the analysis have been offered, the one devised by de Sausure will be adopted.[4] He simply stated that meaning can be communicated because some sign, an utterance or written word, makes a reference, a mental image in the minds of the users, to a referent in the real world. Referents can be particular objects or processes or abstract concepts. Now, language is a conventional system which has evolved to communicate with oneself and others and is predicated, as with all conventional systems, upon the correct use of signs pointing to referents to make the appropriate reference. Here further variables can be identified; the quality of choice of signs by the transmitter and the quality of their understanding by the receiver in the sense of the relation between the sign and the reference as distinguished from the above-mentioned relation between the reference and the referrent. Inaccurate or incorrect use or understanding by either will militate against communication. It should be mentioned that for all the above variables linguistic communication is not completely inexistent, only reduced to partial success.

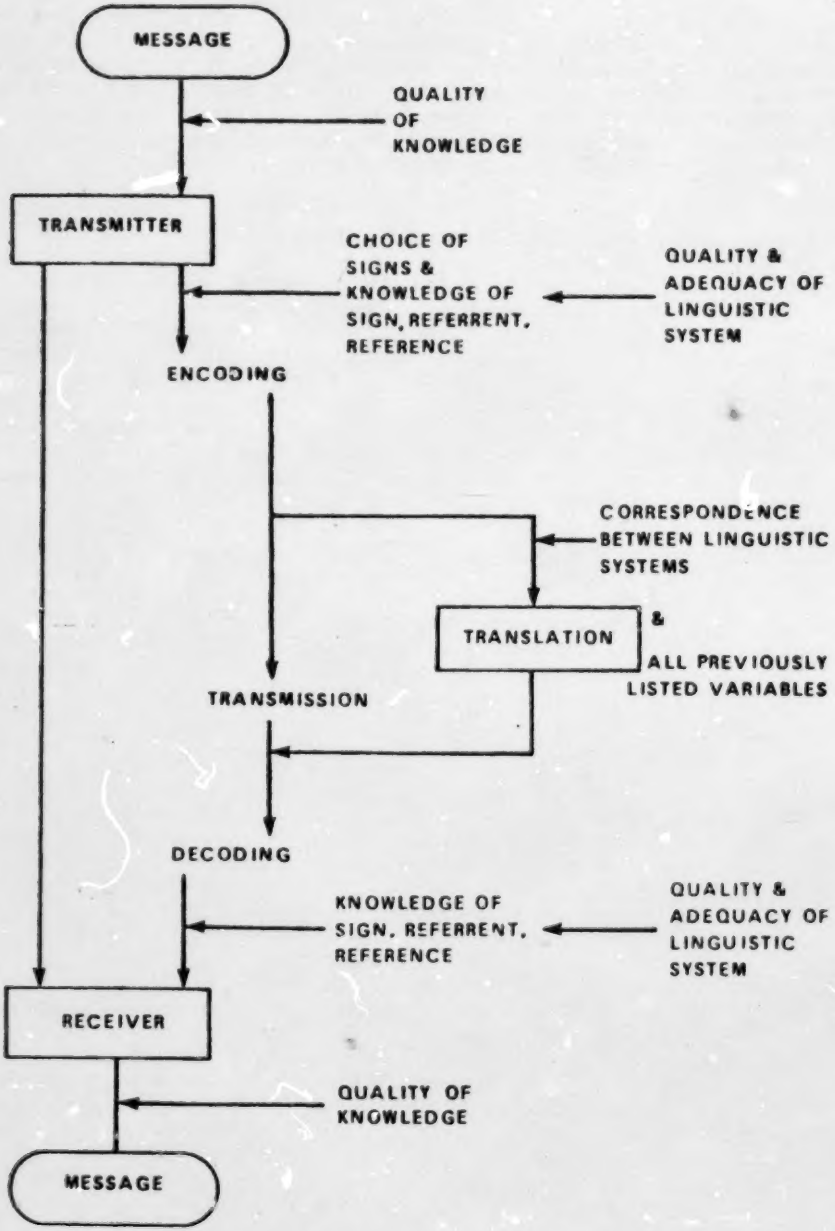
While the subjective variables of quality of use and understanding by members of any linguistic group explain much of the communication process, a series of objective variables must be examined. To this end another triangular model will be introduced. For one referent in an ideal linguistic system there should exist one sign and one definition. That is, there should be a consensus as to a sign pointing to a specific referent and equally a consensus as to the definition of this reference. This ideal system has been described as "one word, one meaning: one meaning, one word" by Borsodi.[5] He makes a very convincing case for its adoption. However, it is far from the existing lexical reality where one referent may be pointed at by many signs while each of these signs might have a number of acceptable definitions. The converse is also true. One word often points at many referents and when used in context with other words to form related expressions the number greatly increases.

One final objective variable of linguistic communication must be added, the adequacy of the language being used. To be adequate it must contain sufficient nouns and pronouns to be able to represent all the universals and particulars both physical and metaphysical. It must contain enough verbs to describe processes and actions, enough adjectives and adverbs to make the necessary modifications and finally enough prefixes, suffixes, conjunctions and prepositions as required by the grammar of the language.

While the discussion to this point has centred on the problems of linguistic communication within one language group, the scenario must be completed by describing the bilingual facet of this process. At first we seem to be faced with the interposition of the same kind of variables as referred to in our discussion on transmitter, encoding, decoding and receiver. The problems, however, are further complicated by the occasional total absence of corresponding symbols for referents in one language or even one idiom. In fact, two linguistic or idiomatic groups perceive the world differently.[6] We are here faced with cultural, geographical and socio-political variables. Eskimos have five different words for snow while Central Africans have never made reference to it. This variable has been distinguished from the subjective quality of an individual's knowledge as it is the whole linguistic group's experience which is absent or unnamed.

The translator who attempts to bridge any inter-linguistic gulf has to take into account all the variables mentioned that are inherent in any attempt to communicate with language plus those various gaps left when looking for a bilingual correspondence. Can it be concluded that because of the number and weight of variables, communication is necessarily doomed to vagueness and ambiguity? Is the sceptic's stance that exact translation is impossible?

From this analysis of the general problem of bilingual communication the JURIVOC research group undertook a project involving the automation of lexicography. While Figure 1 attempts to summarize the general statement of the problem the following sections will develop the framework of analysis of the particular task of Canadian bilingual legal lexicography which will lead to the discussion on methodology. In conclusion, when one uses language for such elevated goals as the communication of truth or the establishment of rights and duties, one must be aware of the tortuous route between sender and receiver and the frailty of the language vehicle. While the exact sciences have developed precise languages of mathematics with which to communicate, law and the social sciences lag considerably in evolving an adequate symbol system to express rights, duties, obligations and responsibilities. Often when a verbal or mathematic explanation are thought inadequate by themselves, scientists use both to permit the receiver to triangulate the messages. Hofeld, disgraced at the linguistic inconsistency of the judges of his time, taught his students the exact definitons of the eight basic legal conceptions and laboured them with exercises on only these fundamentals.[7] More recently individuals have tried to develop algebras of legal relations while the Federal German Republic supplies flow charts and graphic networks to facilitate the implementation of legislation.[8]



Bilingual Communication

Surely the linguistic reality can be faced. It can no longer be presumed that even the trained technicians of the legal system have a "thoroughly good understanding" of what is meant by laws, regulations, orders and procedures. It is to this problem that the JURIVOC group addressed themselves; legal communication. It has been towards its understanding and improvement that the group has researched and developed methods which have, in turn, yielded powerful tools to be employed in the solution of the problem of Canadian bilingual legal communication.

3.1.2 THE PHENOMENON OF LEGAL VOCABULARY

The total set of all utterances concerning law is the phenomenon of the legal vocabulary. Every spoken or written word concerning law in all of time is part of this set. Only a small fraction of this has ever been recorded and even a smaller fraction preserved. This set includes formal laws, judgments and orders as well as informal comments and discussions. While this gives us the complete picture of language usage, in order to grasp the reality of the practical side of this phenomenon an analytic framework proposed by Allen shall be adopted.[9] It is distinct from the analysis of the variables of bilingual communication which discussed individual linguistic message transferring in as far as it focuses on language in general as an objective event. The five dimensions of usage of any language are: modal, situational, cultural, spatial and temporal. The modal dimension is the method or mode of communication either speech or writing. The situational dimension can be represented by a continuum between formality and informality. The cultural dimensions denote the distinctions between standard and non-standard usage varying with factors such as schooling and social contacts. The last two dimensions are those of space and time. The former refers to the localization of an utterance in some linguistic - geographic region while the latter refers to the period of usage, the past usually being denoted as "obsolete" or "archaic".

By using this framework, the types of utterance that are of interest may be classified. The British North American Act, for example, is an extreme of written, formal and standard usage for the Canadian territory between 1867 and today. On the other hand, a comment by a barber to one of his clients about the legality of the Conscription Act in 1939 is verbal, informal and probably non-standard usage. Both of these utterances might theoretically be used by a legal lexicographer. However, the focus of this work is exclusively on legal vocabulary which is written, formal and of standard use in Canada since 1867.

The major legislative sources of legal vocabulary are the statutes and regulations of the ten provincial governments, the federal government and the councils of the larger metropolitan areas. The judicial sources begin at the Supreme Court of Canada then branch to the appeal courts of the provinces and would also encompass the Federal Court and certain specialized tribunals such as labour boards. Other sources which might be considered for analysis are some

doctrine, some procedures and formulae, all codified law and certain standard contracts.

The bilingual sources of vocabulary are surprisingly extensive and all in machine-readable form except the regulations of New Brunswick. This means that a Canadian Bilingual Legal Dictionary could be machine generated from the vocabulary found in the Federal Statutes and Regulations, the Statutes of Quebec and New Brunswick, the Quebec Regulations, the Civil Code, the decisions (since 1970) of the Supreme Court of Canada, and the Federal Court. These cover all the major jurisdictions and form a respectable basis for Canadian bilingual vocabulary use.

3.1.3 LEXICOGRAPHICAL ASSUMPTIONS

Traditionally lexicography has been considered a most arduous and time consuming task essentially because a dictionary maker must collect an enormous quantity of primary data. This process involves selecting word uses from a huge amount of written material, filing the contexts and finally synthesizing the meanings for sometimes hundreds of thousand words and phrases. The initial phase of culling the possible reliable sources for candidates discourages all but the very hardy. It is principally because of this dimension that the trend away from producing general dictionaries has accelerated. The persisting need has been filled by periodically up-dating existing general works and by developing smaller specialized vocabulary works. Terminology data bases which use the computer to store the results of on-going analysis have also appeared. The computer vastly reduces the manipulation required in building and maintaining adequate files; it also simplifies modifications to the data base that result from the discovery of new uses of old terms and the introduction of completely new terms.

The specialized lexicon is usually based on a restricted and well defined corpus of texts. The trend mentioned above has meant that the emphasis is now put on the use of words within a reduced context. This might be called the emergence of the disciplinarian approach to lexicography. By shifting the emphasis to the use of words in special contexts one is confronted with a new methodological problem. It is quite obvious that the core terms of any discipline can be listed, that is, those words and phrases that have been evolved to signal the special concepts of a field; but, more general terms can have different meanings when used in these contexts. The question might be asked as to whether the lexicon of a specialized field includes all the words used when communicating information about that field or just the core terms. The answer to this question forms the basis of our first lexicographical assumption. It states that *any word, despite however general its use might be in the total lexicon of a language, may take on a particular meaning within a specialized context.* This implies that the builders of disciplinarian dictionaries must consider every word as a possible candidate for entry in their final products.

The next assumption states that *the certitude of discovering all the meanings of a word varies with the exhaustiveness of the analysis of the defined corpus of text*. Most terms have a number of major meanings which occur frequently and a series of infrequently used meanings. While our intuition might discern the major uses, it is only by analyzing all occurrences of a word within contexts that one can discover the extent of its use. This approach is all the more valid for bilingual dictionaries. Many authors and many translators have worked on the texts of a specialized discipline and the quality of their work varies greatly. All lexicographers find a spectrum of uses ranging from blatantly erroneous to beautifully polished. Nor is the most frequently used necessarily the best or most correct. The only way to discover the lexicon of a given discipline is to analyse every occurrence of every word of the corpus of its important texts. An important text is defined as one which is formal and has been sanctioned by a legislative, judicial or executive authority.

Prescriptive dictionaries have often been described in opposition to informative dictionaries.10 It seems that there is a qualitatively different type of ponderation required to produce each. The former tries to normalize the use of terms by grading each candidate while the latter exposes, without qualification, all the different discovered uses. The two have been interpreted as different stages in the process of recording language use in the bilingual dictionary format. The informative lexicon is a necessary precursor to the establishment of a prescriptive dictionary. There is but one way of discovering the best uses and translations to be suggested for future usage and that is by first uncovering the full and actual extent of how a term has been used or translated. This simple fact leads to the assumption that states that *competent prescription for a term must be based on the complete set of distinguished uses of that term*. We might generalize this assumption to any lexicographically related product such as a thesaurus, an index or even a reference system. This implies that the process starts empirically by collecting and arranging by some order of precedence all the existing facts of experience, in this case, the actual occurrences of words within their contexts. The parallel activity, more akin to deduction, takes place as the images of the use of each word or group of closely related words starts to form. Patterns and relations are only identified by studying sets of facts.

The next assumption states that *standardisation of lexical entries properly takes place only after each form of the word has been analysed separately*. To facilitate the access to pertinent information the lexicographer may include different tenses of a verb and plurals of nouns under the infinitive and singular forms respectively. Adjectives and adverbs have even been treated this way. The related, broader and narrower term references are other means of permitting access to an entry which indicates the use of the word in all contexts. Because the lexicographical activity has been analysed as an ordering process from heterogeneously distributed vocabulary to a list sorted by some identified order of precedence, it considers the writing of different forms of words and the standardising of dictionary entries as stages in this process. The appropriate time to perform this work is when all distinguished uses and translations have been identified for each word. With general rules to guide the integration, it is essential first to have a complete list of all the components. Hence, some words will

have separate singular and plural noun entries or distinct past and present participle verb entries while others will be grouped under a standardised form.

It has been further assumed that *the lexicographical activity is a process decomposable into discrete phases each involving intellectual and mechanical processing.* The work system approach of systems analysis has been applied to rationalize a symbiotic production system exploiting the best possible contributions from humans and machines. This analysis has yielded the design of a production tool which requires routinized human inputs into a sequence of programmed sorting routines. It has been assumed that by generating mechanically ordered lexical information humans can then more easily create higher quality dictionaries for less time and money.

The assumptions that have been exposed above the basis of strategy for producing bilingual lexicographical aids. When one is concerned with a special vocabulary of a discipline every word must be processed. To be able to compile the complete image of a word's use it is necessary to analyze every occurrence of the word. While developing a dictionary which prescribes correct usuage one must first create the informative lexicon which contains examples of all distinguished uses and translations. Likewise, the grouping and standardising of word forms can only successfully maintain identified semantic nuances if performed after the analysis of the individual words is complete. The way to increase bilingual dictionary production efficiency is to design a symbiotic work system which combines the best of the capabilities of men and machines.

3.1.4 THE CONCEPTS OF COMPUTER USE IN LEXICOGRAPHY

The variety and extent of computer applications to lexicography is unknown. The fields of computational linguistics and mechanical translation have developed recently and with so much vigour that pioneering research is continually being done on a very wide basis.[10] Our own efforts can be seen as a part of this wider effort.

Research and development often resolves itself to solving contingent problems of an intuitively developed approach when faced with a challenge as ambiguously stated as "make a computer produce a bilingual dictionary". Generalized principles and particular rules have yet to be synthesized from the wide variety of experiments carried out. However, there are a series of concepts which, if reflected upon, can inspire creative individuals with an understanding of the problems from a general point of view. Even a random browsing through scientific principles can serve to precipitate an understanding of previously unrecognized facets of any endeavour. Synergy of information is a very powerful tool for discovery and invention.

Artificial intelligence is an important and inspiring concept applicable to computers and the task of building tools to produce lexicographical aids. It recognizes similarities between a type of computer operation and certain human

intellectual processes. This framework highlights an increased work capacity by cooperatively soliciting the help of computers to assume a certain number of routine tasks in a given venture. This "doing more with less" is possible because of the combined effect of working together. The amplification of human power is achieved despite the fact that a computer program can only accomplish a certain percentage of its allotted task. In the area of pattern recognition (in the field of text processing) its success rate is 95%; it fails in 5% of the cases where it has not the stored information or where the context is ambiguous. This does mean, however, that 19 of every 20 examples of a task traditionally thought of as exclusively human has been automated.

The analogy for the development of artificial intelligence can also be found in the training of children and in their later stages of self-learning. A cooperative man-machine undertaking involves communication based upon established conventions and protocol. The initial building of memory, its structuring and programming of access, is initiated and executed by humans. The human must also plan the basic logic routines for a machine to perform any algorithmic analysis or recognition of patterns.

Once the tasks have been defined and the necessary information transferred the better functional abilities of the computer can begin to be used. It has a perfect memory and reading or writing are performed at phenomenal rates. It can retrieve in seconds words and phrases from hundreds of pages of text and it can sort information according to multi-level orders of precedence. It can also work 24 hours a day. These are the more useful functions that can be used to help us reduce the traditionally arduous task of building dictionaries and other lexicographical aids.

The essential concept for computer applications in the field of automated bilingual lexicography is that of symbiotic work systems which cooperatively combine a developing artificial lexicographical intelligence with human synthetic abilities to produce variously structured compilations of word use.

3.5 BILINGUALISM

When processing the vocabulary of a corpus of unilingual texts one is in search of the different uses of words. The end product is usually a dictionary based upon a series of intermediate listings such as key word in contexts, synonyms and related expressions. Vocabulary processing of bilingual texts seemingly requires a whole new perspective. In fact, this is not the case.

It has been assumed that in all possible cases of translation, from literal through adequate or natural to poetic, the messages of the two versions of the text must correspond. That is to say, the different linguistic signs often make dissimilar reference to referents which must correspond. It is this basic equality of referents which makes a process of triangulation possible. The two language

versions form a stable triangle with the shared meaning of the texts. Now, this same triangulation is fundamental to the process of compiling a unilingual dictionary. We say that two occurrences of a word in their contexts have the same use if, despite different signs, they both point to the same referent. Another example from a totally different context but still dealing with communication is that many scientists will use both verbal and mathematical descriptions of a phenomena to allow a reader the greatest possibility of capturing the meaning by what is called triangulation of references.

This analysis has lead the authors to expand the applicability of the word "use" to include translations. The lexical data files used for automatic analysis store related expressions and translations of a key word while the output lexicons or dictionaries include different translations of a key word, standardised forms of related expressions and their translations as well as examples of distinguished meanings in context. The use of an English key word in bilingual texts, as reflected by the final entries in a dictionary, is a linked pair of French-English correspondents. Differentiation of use is done not only by a pure alphabetic sort but also by contextual meaning differences located in either of the versions. This expanded definition of "use" has allowed us to process bilingual texts much in the same way as one would approach unilingual vocabulary analysis but with the added assurance of being able to capture the meanings by having two ways (i.e. French and English) of stating the same concept. This has greatly simplified the automation of the task of producing bilingual dictionaries.

1. *Random House Dictionary, Unabridged Edition*, New York, Random House, 1971.

2. SHANNON, C. and WEAVER, W., *Mathematical Theory of Information*, University of Illinois Press, 1962.

3. BLOOMFIELD, L., GARDINER, A.H., LYONS, J., MORRIS, C., OGDEN, C., de SAUSSURE, F., *et al.*

4. de SAUSSURE, F., *Cours de linguistique générale*, Paris, Payot, 1955.

5. BORSODI, R., *The Definition of Definition*, Boston, Porter Sargent, 1967, p. 14

6. WHORF, Benjamin L., SAPER, Edward, von HUMBOLDT, Wilhelm, *et al*

7. HOHFELD, W. N., *Fundamental Legal Conceptions*, Yale University Press, 1919, p.23

8. DICKEY, Anthony, *A Fresh Approach to the Analysis of Legal Relations*, McGill Law Journal, 1974 (20) p. 261.

9. ALLEN, Harold B., *Readings in Applied English Linguistics*, New York, 1964, p. 273.

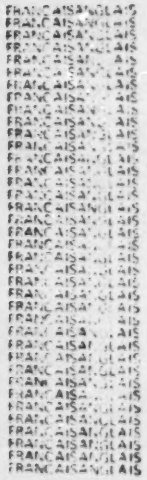
10. WELLS, Ronald A., *Dictionaries and the Authoritarian Tradition*, Mouton, The Hague, 1973, p. 87.

JURIVOC

WORDS COMPUTERS AND COMMUNICATION IN LAW

Viateur Bergeron
David C. Burke

Jeanne Leclerc-Houde
Michael Mepham
Roland Serrat



LEXICOGRAPHIE BILINGUISME JURIDIQUE ET ORDINATEUR

16

JURIVOC

Les Éditions de l'Université d'Ottawa